CLAIMS

What is claimed is

1. A method, comprising determining, in a centralized fashion, paths for flows within a multi-stage network made up of clusters of processing nodes, and encoding node selection information representing flow path decisions for all clusters of the multi-stage network in packets of each flow within the multi-stage network.

2. The method of claim 1, wherein the paths are determined without knowledge of whether or not packets of a particular flow will actually traverse specific ones of the clusters within the multi-stage network.

3. The method of claim 2, wherein the node selection information is encoded in the packets by replacing other header information in the packets with the node selection information.

4. The method of claim 3, wherein original header information present prior to encoding of the node selection information is restored to the packets prior to packet processing by application programs executing at nodes of the multi-stage network.

5. The method of claim 2, wherein the node selection information is encoded in the packets by appending the node selection information to the packets.

6. The method of claim 5, wherein the node selection information is stripped prior to packet processing by application programs executing at nodes of the multi-stage network.

7. The method of claim 2, wherein individual nodes of each cluster are selected for inclusion in the paths according to load balancing criteria.

8. The method of claim 7, wherein the load balancing criteria comprises a round-robin selection methodology.

9. The method of claim 7, wherein the load balancing criteria comprises a weighted round-robin selection methodology.

10. The method of claim 7, wherein the load balancing criteria comprises real time information from the nodes of the multi-stage network as to current load conditions.

11. The method of claim 2, wherein prior to determining the paths for the flows, the flows are classified using a set of information present in the packets and according to criteria established by a network administrator, and flow paths are assigned based on policies set by the network administrator.

12. The method of claim 1, further comprising determining actual flow routes on a node-by-node basis as packets of the flows traverse the flow paths within the multi-stage network.

13. The method of claim 12, wherein the actual flow routes do not include all of the clusters specified in the flow paths.

14. The method of claim 13, wherein information regarding the actual flow routes is used in determining subsequent new flow paths.

15. A method, comprising determining, in a distributed fashion, paths for flows within a multi-stage network made up of clusters of processing nodes, and encoding node selection information representing flow path decisions in packets of each flow within the multi-stage network.

16. The method of claim 15, wherein the node selection information is encoded in the packets by replacing other header information in the packets with the node selection information.

17. The method of claim 15, wherein original header information present prior to encoding of the node selection information is restored to packet headers prior to packet processing by application programs executing at the nodes of the multi-stage network.

18. The method of claim 15, wherein node selection is made on the basis of applying a hash function to a flow identifier encoded within the packet.

19. The method of claim 18, wherein the node selection information is appended to each packet.

20. The method of claim 19, wherein the node selection information is stripped prior to packet processing by application programs executing at the nodes of the multi-stage network.

21. The method of claim 20, wherein the nodes are grouped into various clusters, and individual nodes of each cluster are selected according to load balancing criteria.

22. The method of claim 21, wherein the load balancing criteria comprises a round-robin selection methodology.

23. A method, comprising replicating, at a node of a multi-stage network in which nodes are grouped into clusters of nodes having similar application functionality, an entire flow, in both directions, and designating a flow path for a resulting replicated flow that is different than an original flow path for an original flow from which the replicated flow was produced.

24. The method of claim 23, wherein the replicated flow is terminated at a terminating node that receives both the replicated flow and the original flow.

25. The method of claim 24, wherein the terminating node records the contents of the replicated flow and the original flow for analysis

26. A system comprising a virtual connectivity grid overlaid on a physical network in which nodes are coupled to one or more physical switches through respective

physical interfaces, the virtual connectivity grid including virtual interfaces overlaid over the node physical interfaces so as to provide communicative coupling of the nodes to one or more virtual networks established within the virtual connectivity grid, the communicative coupling being established by virtual links arranged so as to emulate physical connections in a desired connectivity pattern.

27. The system of claim 26, wherein each of the nodes includes a flow management module configured to manipulate packets transmitted within the virtual connectivity grid by modifying information contained in the packets to reflect flow management decisions.

28. The system of claim 27, wherein the flow management decisions are made in a centralized fashion and on a flow-by-flow basis within the virtual connectivity grid at those of the nodes which are coupled to one or more external networks and which receive flows from the one or more external networks.

29. The system of claim 28, wherein the packets are manipulated by appending information specifying flow routing decisions.

30. The system of claim 28, wherein the packets are manipulated by replacing previously encoded information within the packet with new information specifying flow routing decisions.

31. The system of claim 28, wherein the packets are manipulated by adding network topology state information to the packets, the network topology state information reflecting the communicative coupling of the nodes.

32. The system of claim 28, wherein the packets are manipulated by adding authentication information.

33. The system of claim 28, wherein flow path decisions are mode without knowledge of which of the nodes a particular flow will actually traverse.

34. The system of claim 26, wherein the system is implemented using a plurality of stand-alone general purpose and/or special purpose computers having software stored therein as the nodes, and local area network (LAN) switches that support virtual local area networks (VLANs) and jumbo frames as the one or more physical switches.

35. The system of claim 34, wherein the software comprises flow management software configured to cause the computers to adaptively route data flows through a multi-stage network defined by the virtual connectivity grid and wherein each stage of the multi-stage network comprises a cluster of similarly configured applications performing one or more network services.

36. The system of claim 34, wherein at least one of the computers stores resource management and provisioning software and remaining ones of the computers store provisioning agent software that is configured to cause the remaining ones of the

computers to adaptively provision a required capacity of one or more network services as driven by application demand, network performance or application response time.

37. The system of claim 36, wherein provisioning is performed via human interaction through a graphical user interface using icons that represent the nodes and the virtual interfaces and creating interconnections between them.

38. The system of claim 36, wherein provisioning is performed via human interaction through a programmatic interface using data records that provide attributes of the nodes and the virtual interfaces and specifying desired interconnections between them.

39. The system of claim 26, wherein the virtual connectivity grid is implemented using one or more blade servers, each blade of the blade servers being one of the nodes and storing software.

40. The system of claim 39, wherein the software comprises flow management software configured to cause processors of the blades to adaptively route data flows through a multi-stage network defined by the virtual connectivity grid and wherein each stage of the multi-stage network comprises a cluster of similarly configured applications performing one or more network services

41. The system of claim 39, wherein at least one of the blades stores resource management and provisioning software and remaining ones of the blades store

provisioning agent software that is configured to cause processors of the remaining ones of the blades to adaptively provision a required capacity of one or more network services as driven by application demand, network performance or application response time.

42. The system of claim 41, wherein provisioning is performed via human interaction through a graphical user interface using icons that represent the nodes and the virtual interfaces and creating interconnections between them.

43. The system of claim 41, wherein provisioning is performed via human interaction through a programmatic interface using data records that provide attributes of the nodes and the virtual interfaces and specifying desired interconnections between them.

44. The system of claim 26 wherein the virtual connectivity grid is implemented using one or more blade servers and one or more stand-alone general purpose and/or special purpose computers as the nodes and one or more local area network (LAN) switches that support virtual local area networks (VLANs) and jumbo frames, external to the blade servers, as the one or more physical switches.

45. The system of claim 44, wherein each of the blades and the computers stores flow management software configured to cause processors of the blades and the computers to adaptively route data flows through a multi-stage network defined by the virtual connectivity grid and wherein each stage of the multi-stage network

comprises a cluster of similarly configured applications performing one or more network services

46. The system of claim 44, wherein at least one of the computers or the blades stores resource management and provisioning software and remaining ones of the computers and/or the blades store provisioning agent software that is configured to cause the remaining ones of the computers and/or processors of the blades to adaptively provision a required capacity of one or more network services as driven by application demand, network performance or application response time

47. The system of claim 44, wherein provisioning is performed via human interaction through a graphical user interface using icons that represent the nodes and the virtual interfaces and creating interconnections between them.

48. The system of claim 44, wherein provisioning is performed via human interaction through a programmatic interface using data records that provide attributes of the nodes and the virtual interfaces and specifying desired interconnections between them.

49. A method, comprising establishing a virtual connectivity grid configured to permit arbitrary interconnections among a first number of computer systems within a computer network, each of the computer systems being communicatively coupled to respective ports of one or more physical network switching devices, through a second number of virtual links that emulate physical network connectivity

mechanisms as a result of configurations of one or more virtual networks (VLANs) overlaid on ports of the physical network switching devices.

50. The method of claim 49, wherein at least one of the virtual links emulates, in a fully switched environment, a cable that provides point-to-point connectivity between two of the computer systems.

51. The method of claim 50, wherein the point-to-point connectivity is unidirectional.

52. The method of claim 50, wherein the point-to-point connectivity is bidirectional.

53. The method of claim 49, wherein at least one of the virtual links emulates, in a fully switched environment, a cable that provides point-to-multipoint connectivity between a root one of the computer systems and a plurality of leaf ones of the computer systems.

54. The method of claim 53, wherein the point-to-multipoint connectivity is unidirectional.

55. The method of claim 54, wherein the point-to-multipoint connectivity is bidirectional.

56. The method of claim 49, wherein at least one of the virtual links emulates, in a fully switched environment, a hub that provides multi-access, broadcast capable

connectivity to a plurality of the computer systems, each operating in either a promiscuous or non-promiscuous mode.

57. The method of claim 49, wherein at least one of the virtual links emulates, in a fully switched environment, a switch that provides multi-access, broadcast capable connectivity to a plurality of the computer systems while providing unicast traffic isolation.

58. The method of claim 49, wherein at least one of the virtual links emulates a transparently bridged connection between a network external to the virtual connectivity grid and a plurality of the computer systems, the bridged connection providing data path transparency for interactions between devices within the network external to the virtual connectivity grid and the plurality of the computer systems.

59. The method of claim 49, wherein prior to transmission across at least one of the virtual links a packet is modified so as to have its original header information contained therein appended to an end of the packet.

60. The method of claim 59, wherein prior to transmission across the at least one of the virtual links the packet is further modified so as to have new header information substituted for the original header information.

61. The method of claim 60, wherein the new header information includes flow path decision information for the packet within the virtual connectivity grid.

62. The method of claim 59, wherein authentication information is appended to the packet, the authentication information serving to identify which of the computer systems is transmitting the packet.

63. The method of claim 49, wherein upon reception at one of the computer systems a packet has its original header information restored with original header information for the packet that is appended to an end of the packet.